18. Explain the different types of Machine Learning algorithms with examples.

19. Discuss in detail the Hadoop framework, Spark.

20. Write a detailed note on disease prediction case study in Data Science.

———————

NOVEMBER/DECEMBER 2025

23UECS53A — INTRODUCTION TO DATA SCIENCE

Time : Three hours Maximum : 75 marks

SECTION A — (10 × 2 = 20 marks)

Answer ALL questions.

Shanmuga Industries Arts & Science College T.V.Malai

1. Define Data Science.

2. What is Big Data ecosystem?

3. What are research goals in Data Science?

4. Define data reduction.

5. What is supervised learning?

6. Give two examples of unsupervised learning algorithms.

7. What is Hadoop framework?

8. Expand ACID and CAP in the context of databases.

9. What is disease profiling?

10. Write any two steps involved in disease prediction using Data Science.

SECTION B — (5 × 5 = 25 marks)

Answer ALL questions.

11. (a) Explain the benefits and uses of Data Science.

Or

(b) Describe the role of Big Data ecosystem in Data Science.

12. (a) Explain the steps involved in the EDA process with an example.

Or

(b) Write short notes on Data Retrieval and Data Transformation.

13. (a) Differentiate between Supervised and Unsupervised learning with examples.

Or

(b) Explain Semi-supervised learning with a suitable example.

14. (a) Discuss the components of the Hadoop framework.

Or

(b) Explain Spark and its advantages over MapReduce.

15. (a) Explain the steps involved in prediction of disease using Data Science.

Or

(b) Write short notes on Data preparation and Disease profiling.

Shanmuga Industries Arts & Science College T.V.Malai

SECTION C — (3 × 10 = 30 marks)

Answer any THREE questions.

16. Explain in detail the concept of bigdata ecosystem and Data Science.

17. Describe the complete Data Science process with neat diagram.